Stochastic Mechanics
Random Media
Signal Processing and Image Synthesis
Mathematical Economics and Finance
Stochastic Optimization
Stochastic Control

Applications of Mathematics

Stochastic Modelling and Applied Probability

31

Edited by
I. Karatzas
M. Yor

Advisory Board
P. Brémaud
E. Carlen
W. Fleming
D. Geman
G. Grimmett
G. Papanicolaou
J. Scheinkman

Springer
Science+Business Media, LLC

Applications of Mathematics

1 Fleming/Rishel, **Deterministic and Stochastic Optimal Control** (1975)
2 Marchuk, **Methods of Numerical Mathematics,** Second Ed. (1982)
3 Balakrishnan, **Applied Functional Analysis,** Second Ed. (1981)
4 Borovkov, **Stochastic Processes in Queueing Theory** (1976)
5 Liptser/Shiryayev, **Statistics of Random Processes I: General Theory,** Second Ed. (1977)
6 Liptser/Shiryayev, **Statistics of Random Processes II: Applications,** Second Ed. (1978)
7 Vorob'ev, **Game Theory: Lectures for Economists and Systems Scientists** (1977)
8 Shiryayev, **Optimal Stopping Rules** (1978)
9 Ibragimov/Rozanov, **Gaussian Random Processes** (1978)
10 Wonham, **Linear Multivariable Control: A Geometric Approach,** Third Ed. (1985)
11 Hida, **Brownian Motion** (1980)
12 Hestenes, **Conjugate Direction Methods in Optimization** (1980)
13 Kallianpur, **Stochastic Filtering Theory** (1980)
14 Krylov, **Controlled Diffusion Processes** (1980)
15 Prabhu, **Stochastic Storage Processes: Queues, Insurance Risk, Dams, and Data Communication,** Second Ed. (1998)
16 Ibragimov/Has'minskii, **Statistical Estimation: Asymptotic Theory** (1981)
17 Cesari, **Optimization: Theory and Applications** (1982)
18 Elliott, **Stochastic Calculus and Applications** (1982)
19 Marchuk/Shaidourov, **Difference Methods and Their Extrapolations** (1983)
20 Hijab, **Stabilization of Control Systems** (1986)
21 Protter, **Stochastic Integration and Differential Equations** (1990)
22 Benveniste/Métivier/Priouret, **Adaptive Algorithms and Stochastic Approximations** (1990)
23 Kloeden/Platen, **Numerical Solution of Stochastic Differential Equations** (1992)
24 Kushner/Dupuis, **Numerical Methods for Stochastic Control Problems in Continuous Time,** Second Ed. (2001)
25 Fleming/Soner, **Controlled Markov Processes and Viscosity Solutions** (1993)
26 Baccelli/Brémaud, **Elements of Queueing Theory** (1994)
27 Winkler, **Image Analysis, Random Fields, and Dynamic Monte Carlo Methods: An Introduction to Mathematical Aspects** (1994)
28 Kalpazidou, **Cycle Representations of Markov Processes** (1995)
29 Elliott/Aggoun/Moore, **Hidden Markov Models: Estimation and Control** (1995)
30 Hernández-Lerma/Lasserre, **Discrete-Time Markov Control Processes: Basic Optimality Criteria** (1996)
31 Devroye/Györfi/Lugosi, **A Probabilistic Theory of Pattern Recognition** (1996)
32 Maitra/Sudderth, **Discrete Gambling and Stochastic Games** (1996)

(continued after index)

Luc Devroye László Györfi
Gábor Lugosi

A Probabilistic Theory of Pattern Recognition

With 99 Figures

Springer

Luc Devroye
School of Computer Science
McGill University
Montreal, Quebec, H3A 2A7
Canada

László Györfi
Gábor Lugosi
Department of Mathematics and
Computer Science
Technical University of Budapest
Budapest
Hungary

Mathematics Subject Classification (1991): 68T10, 68T05, 62G07, 62H30

Library of Congress Cataloging-in-Publication Data
Devroye, Luc.
A probabilistic theory of pattern recognition/Luc Devroye,
László Györfi, Gábor Lugosi.
p. cm.
Includes bibliographical references and index.

1. Pattern perception. 2. Probabilities. I. Györfi, László.
II. Lugosi, Gábor. III. Title.
Q327.D5 1996
003'.52'015192—dc20 95-44633

Printed on acid-free paper.

Originally published by Springer-Verlag New York, Inc. in 1996
Softcover reprint of the hardcover 1st edition 1996

Production managed by Francine McNeill; manufacturing supervised by Jeffrey Taub.
Photocomposed copy prepared using Springer's svsing.sty macro.
9 8 7 6 5 4 3

SPIN 10830936

ISBN 978-1-4612-6877-2 ISBN 978-1-4612-0711-5 (eBook)
DOI 10.1007/978-1-4612-0711-5

Preface

Life is just a long random walk. Things are created because the circumstances happen to be right. More often than not, creations, such as this book, are accidental. Nonparametric estimation came to life in the fifties and sixties and started developing at a frenzied pace in the late sixties, engulfing pattern recognition in its growth. In the mid-sixties, two young men, Tom Cover and Peter Hart, showed the world that the nearest neighbor rule in all its simplicity was guaranteed to err at most twice as often as the best possible discrimination method. Tom's results had a profound influence on Terry Wagner, who became a Professor at the University of Texas at Austin and brought probabilistic rigor to the young field of nonparametric estimation. Around 1971, Vapnik and Chervonenkis started publishing a revolutionary series of papers with deep implications in pattern recognition, but their work was not well known at the time. However, Tom and Terry had noticed the potential of the work, and Terry asked Luc Devroye to read that work in preparation for his Ph.D. dissertation at the University of Texas. The year was 1974. Luc ended up in Texas quite by accident thanks to a tip by his friend and fellow Belgian Willy Wouters, who matched him up with Terry. By the time Luc's dissertation was published in 1976, pattern recognition had taken off in earnest. On the theoretical side, important properties were still being discovered. In 1977, Stone stunned the nonparametric community by showing that there are nonparametric rules that are convergent for all distributions of the data. This is called distribution-free or universal consistency, and it is what makes nonparametric methods so attractive. Yet, very few researchers were concerned with universal consistency—one notable exception was Laci Györfi, who at that time worked in Budapest amid an energetic group of nonparametric specialists that included Sándor Csibi, József Fritz, and Pál Révész.

So, linked by a common vision, Luc and Laci decided to join forces in the early eighties. In 1982, they wrote six chapters of a book on nonparametric regression function estimation, but these were never published. In fact, the notes are still in drawers in their offices today. They felt that the subject had not matured yet. A book on nonparametric density estimation saw the light in 1985. Unfortunately, as true baby-boomers, neither Luc nor Laci had the time after 1985 to write a text on nonparametric pattern recognition. Enter Gábor Lugosi, who obtained his doctoral degree under Laci's supervision in 1991. Gábor had prepared a set of rough course notes on the subject around 1992 and proposed to coordinate the project—this book—in 1993. With renewed energy, we set out to write the book that we should have written at least ten years ago. Discussions and work sessions were held in Budapest, Montreal, Leuven, and Louvain-La-Neuve. In Leuven, our gracious hosts were Ed van der Meulen and Jan Beirlant, and in Louvain-La-Neuve, we were gastronomically and spiritually supported by Léopold Simar and Irène Gijbels. We thank all of them. New results accumulated, and we had to resist the temptation to publish these in journals. Finally, in May 1995, the manuscript had bloated to such extent that it had to be sent to the publisher, for otherwise it would have become an encyclopedia. Some important unanswered questions were quickly turned into masochistic exercises or wild conjectures. We will explain subject selection, classroom use, chapter dependence, and personal viewpoints in the Introduction. We do apologize, of course, for all remaining errors.

We were touched, influenced, guided, and taught by many people. Terry Wagner's rigor and taste for beautiful nonparametric problems have infected us for life. We thank our past and present coauthors on nonparametric papers, Alain Berlinet, Michel Broniatowski, Ricardo Cao, Paul Deheuvels, András Faragó, Adam Krzyżak, Tamás Linder, Andrew Nobel, Mirek Pawlak, Igor Vajda, Harro Walk, and Ken Zeger. Tamás Linder read most of the book and provided invaluable feedback. His help is especially appreciated. Several chapters were critically read by students in Budapest. We thank all of them, especially András Antos, Miklós Csűrös, Balázs Kégl, István Páli, and Márti Pintér. Finally, here is an alphabetically ordered list of friends who directly or indirectly contributed to our knowledge and love of nonparametrics: Andrew and Roger Barron, Denis Bosq, Prabhir Burman, Tom Cover, Antonio Cuevas, Pierre Devijver, Ricardo Fraiman, Ned Glick, Wenceslao Gonzalez-Manteiga, Peter Hall, Eiichi Isogai, Ed Mack, Arthur Nádas, Georg Pflug, George Roussas, Winfried Stute, Tamás Szabados, Godfried Toussaint, Sid Yakowitz, and Yannis Yatracos.

Gábor diligently typed the entire manuscript and coordinated all contributions. He became quite a TEXpert in the process. Several figures were made by `idraw` and `xfig` by Gábor and Luc. Most of the drawings were directly programmed in PostScript by Luc and an undergraduate student at McGill University, Hisham Petry, to whom we are grateful. For Gábor, this book comes at the beginning of his career. Unfortunately, the other two authors are not so lucky. As both Luc and Laci felt that they would probably not write another book on nonparametric pattern recognition—the random walk must go on—they decided to put their general

view of the subject area on paper while trying to separate the important from the irrelevant. Surely, this has contributed to the length of the text.

So far, our random excursions have been happy ones. Coincidentally, Luc is married to Bea, the most understanding woman in the world, and happens to have two great daughters, Natasha and Birgit, who do not stray off their random courses. Similarly, Laci has an equally wonderful wife, Kati, and two children with steady compasses, Kati and János. During the preparations of this book, Gábor met a wonderful girl, Arrate. They have recently decided to tie their lives together.

On the less amorous and glamorous side, we gratefully acknowledge the research support of NSERC CANADA, FCAR QUEBEC, OTKA HUNGARY, and the exchange program between the Hungarian Academy of Sciences and the Royal Belgian Academy of Sciences. Early versions of this text were tried out in some classes at the Technical University of Budapest, Katholieke Universiteit Leuven, Universität Stuttgart, and Université Montpellier II. We would like to thank those students for their help in making this a better book.

Montreal, Quebec, Canada — Luc Devroye
Budapest, Hungary — Laci Györfi
Budapest, Hungary — Gábor Lugosi

Contents